

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-303528

(43)Date of publication of application : 16.11.1993

(51)Int.Cl.

G06F 12/08

(21)Application number : 04-134323

(71)Applicant : OKI ELECTRIC IND CO LTD

(22)Date of filing : 27.04.1992

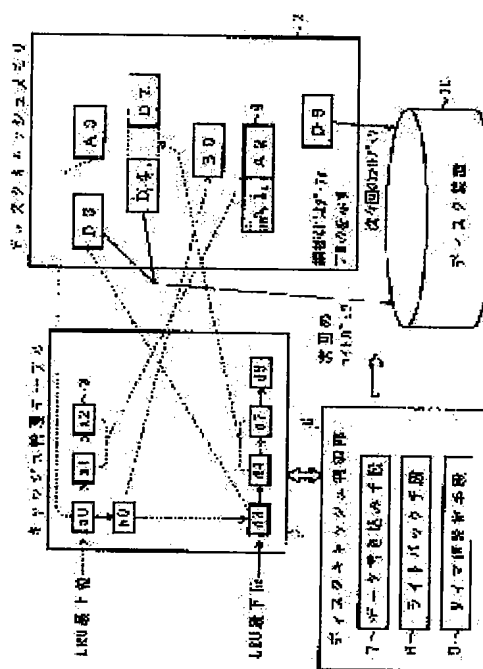
(72)Inventor : YASUDA HIROYUKI

(54) WRITE-BACK DISK CACHE DEVICE

(57)Abstract:

PURPOSE: To decrease the frequency of write-back operation and improve the throughput.

CONSTITUTION: A pointer is provided so as to link individual control blocks 3 in a cache control table 1 in the order of the addresses in an area on a disk device 10. Then when the linked block group includes a dirty block, only the LRU pointer of the head block in the group is made valid. In the write-back operation, plural blocks among blocks belonging to the lowest order in the LRU in the group including the dirty block are written at a time as much as possible in order from the block of the low-order address. Further, a timer which generates the timing for the start of next write-back operation after the write or write-back operation is provided. Its timer value is determined depending upon the number of dirty blocks left in a disk cache memory 2 at the point of time.



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平5-303528

(43)公開日 平成 5 年(1993)11月16日

(51)Int.Cl.⁵
G 0 6 F 12/08

識別記号
3 2 0

庁内整理番号
7232-5B

F I

技術表示箇所

審査請求 未請求 請求項の数 2 (全 16 頁)

(21)出願番号 特願平4-134323

(22)出願日 平成 4 年(1992) 4 月27日

(71)出願人 000000295

沖電気工業株式会社
東京都港区虎ノ門 1 丁目 7 番12号

(72)発明者 保田 浩之

東京都港区虎ノ門 1 丁目 7 番12号 沖電気
工業株式会社内

(74)代理人 弁理士 佐藤 幸男

(54)【発明の名称】 ライトバック式ディスクキャッシュ装置

(57)【要約】

【目的】 ライトバックの回数を減少させ、スループットを向上させる。

【構成】 キャッシュ管理テーブル 1 上で個々の管理ブロック 3 をディスク装置 1 0 上の領域内でのアドレス順にリンクするためのポインタを設ける。そして、リンクされたブロック群にダーティブロックが含まれる場合はそのグループ内の先頭ブロックの L R U ポインタのみを有効とする。ライトバック時にはダーティブロックを含むグループ内で L R U で最も下位に属するものの中で、下位アドレスのブロックから順にできる限り複数ブロックを一括して書き出す。また、ライト又はライトバック終了後、次にライトバックを起動するタイミングを発生させるタイマを設ける。そのタイマ値はその時点でディスクキャッシュメモリ 2 上に残っているダーティブロック数によって決定する。

【特許請求の範囲】

【請求項１】 ディスク装置上に格納されたデータをブロック毎に転記するディスクキャッシュメモリと、前記ディスク装置へのデータの書き込みの発生時は、当該ディスクキャッシュメモリに転記されている該当するブロック内のデータのみを更新するデータ書き込み手段と、前記ディスク装置上の連続する１トラック分程度の複数のブロックを１つのグループとし、当該グループ内のブロックを前記ディスク装置上の同一トラック内のアドレス順にリンクして管理するキャッシュ管理テーブルと、当該キャッシュ管理テーブル上に前記ディスクキャッシュメモリ上でのみ更新されているブロックを含むグループがあるときは、当該グループ内の連続する複数のブロックを１回のアクセスにより前記ディスク装置上に書き戻すライトバック手段とから成ることを特徴とするライトバック式ディスクキャッシュ装置。

【請求項２】 ディスク装置上に格納されたデータをブロック毎に転記するディスクキャッシュメモリと、前記ディスク装置へのデータの書き込みの発生時は、当該ディスクキャッシュメモリに転記されているブロック内のデータのみを更新するデータ書き込み手段と、前記キャッシュメモリ上でのみ更新されているブロックを前記ディスク装置上に書き戻すライトバック手段と、当該ライトバック手段の動作が発生する時間間隔を決定するタイマ値を、前記ディスクキャッシュメモリ上でのみ更新されているブロックの数が多くなるに従って小さく設定するタイマ値設定手段とから成ることを特徴とするライトバック式ディスクキャッシュ装置。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】 本発明は、ディスクキャッシュメモリの管理とライトバックのタイミング制御を行なうライトバック式ディスクキャッシュ装置に関するものである。

【０００２】

【従来の技術】 ディスクキャッシュ装置は、磁気ディスク装置に対するアクセス時間を短縮するために、主記憶装置に対するキャッシュメモリの原理を利用して構成されたものである。

【０００３】 図２及び図３は、従来のライトバック式ディスクキャッシュ装置のブロック図であり、これらの図はそれぞれリード時及びライト時の同装置の動作を説明するものである。図示の装置は、システムバスに接続された、中央処理装置２１と、主記憶装置２２と、ディスクキャッシュ装置２５等から成る。中央処理装置２１は、磁気ディスク装置２６に格納されたプログラムやデータを主記憶装置２２に読み出してデータ処理を行なう。主記憶装置２２は、中央処理装置２１で実行されるプログラムや作業中のデータを一時的に格納する。ディ

スクキャッシュ装置２５は、ディスクキャッシュ制御部２３と、ディスクキャッシュメモリ部２８とから成る。ディスクキャッシュ制御部２３は、磁気ディスク装置２６とディスクキャッシュメモリ部２８との間のデータ転送を制御する。ディスクキャッシュメモリ部２８は、ＲＡＭから成り、磁気ディスク装置２６の最小アクセス単位であるブロック毎にデータを格納する。磁気ディスク装置２６は、ハードディスク又はフロッピディスクから成る。尚、磁気ディスクに限らず、光ディスクでも同様である。

【０００４】 図２は、中央処理装置２１から磁気ディスク装置２６上の領域２７データのリード要求を受けたときの動作を示したものである。このとき、ディスクキャッシュ装置２５は、まず、ディスクキャッシュメモリ部２４上に該当する領域のデータの写しがあるかどうかを検索する。ディスクキャッシュメモリ部２４上にそのデータが存在しなかった場合をリードミスと呼ぶ。この場合、磁気ディスク装置２６から領域２７のデータを読み出して主記憶装置２２上の指定された領域２９へ転送すると同時に、ディスクキャッシュメモリ部２４上の領域２８に格納する。この後、同一データのリード要求を受けた場合、ディスクキャッシュメモリ部２４上にそのデータがまだ存在していた場合がリードヒットである。このときは磁気ディスク装置２６をアクセスせずにディスクキャッシュメモリ部２４上の領域２８から直接主記憶装置２２に転送する。このため、リードヒット時にはアクセス時間を大幅に短縮することができる。

【０００５】 図３は、ＣＰＵ２１から主記憶装置２２上の領域３０のデータの、磁気ディスク装置２６上の領域３２へのライト要求を受けたときの動作を示したものである。ライトバック方式のディスクキャッシュ装置２５はこのときも、ディスクキャッシュメモリ部２４上に磁気ディスク装置２６上の領域３２のデータの写しの有無を検索して、それが存在した場合はディスクキャッシュメモリ部２４上の該当する領域にライトデータを上書きし、そうでなければ領域を新規に割当ててそこに格納する。こうして、ディスクキャッシュメモリ部２４上の領域３１へのデータの転送を終えた時点でＣＰＵ２１へライト終了を通知する。その後、適当な時点で、ディスクキャッシュ装置２５は、領域３１のデータを磁気ディスク装置２６の領域３２へライトする。この動作をライトバックと呼ぶ。これにより、ライト時にはデータを格納する領域の割当てが可能な限り、アクセスタイムを見かけ上短縮することができるのである。ディスクキャッシュメモリ上のデータの管理単位をブロック、その大きさをブロックサイズと呼ぶ。ディスクキャッシュ装置では磁気ディスク装置へのアクセスの局所性を利用してヒット率をたかめるために、ディスクキャッシュメモリ内の各ブロックは、最近にアクセスされたもののほどキャッシュ上に常駐する優先度を高くする、いわゆる“Least

Recently Used (以下、「LRU」という)の原理で管理される。

【0006】図4は、従来のライトバックキャッシュの管理方式の説明図である。ディスクキャッシュ装置はディスクキャッシュメモリ42を管理するためのキャッシュ管理テーブル41を持つ。そして、そこに、ディスクキャッシュメモリ42上のキャッシュブロック44と一体に対応する管理ブロック43を格納している。各管理ブロック43は、対応する管理ブロック43内のLRUポインタというポインタによってLRU順にリンクされている。そして、これと、LRUの最上位と最下位を指し示すポインタによりLRU管理が実現されている。図4では、このポインタによるリンクの様子を矢印で模式的に表わしている。この図で、キャッシュ管理テーブル41上の管理ブロック43の順序は、物理的な配置とは関係しない論理的なものであり、LRUポインタによるリンクをつなぎかえるだけで順序を容易に変更することができる。

【0007】ライトバック方式のディスクキャッシュ装置では、磁気ディスク装置にライトバックする前のライトデータを含むキャッシュブロックがディスクキャッシュメモリ42上に存在する。これをダーティブロックという。ディスクキャッシュ装置は、ライトによるデータの更新を磁気ディスク装置に反映させるために、適当なタイミングでLRU最下位に最も近いダーティブロックを選択してライトバックする。図4の場合では、次にライトバックの対象となるのはブロックDである。

【0008】図5は、従来のライト時のキャッシュ制御を説明するフローチャートである。ここで、CPUからのライト要求を受けたときのキャッシュ制御の例を図5のフローチャートを基に説明する。ディスクキャッシュ装置は、まずキャッシュ管理テーブルを検索して磁気ディスク装置上のライト対象となる領域のデータを格納したブロックがディスクキャッシュメモリ上に存在するか否かを判定する(ステップS21、S22)。そして、存在した場合はそのブロックにデータを上書きする(ステップS30)。そうでなかった場合は、次にディスクキャッシュメモリ上の空きブロックの有無を確認する(ステップS23)。もし空きブロックがなければ、LRU最下位のブロックを追いついてそのブロックを空きブロックにする必要がある(ステップS24、S25、S26)。そのブロックがダーティブロックでなければ単に無効化するのみでよく(ステップS24、S26)、そうでなければそのブロックのデータをライトバックした後に無効化して空きブロックとする(ステップS24、S25、S26)。そして、空きブロックから新たにデータを割当て、そこにデータをライトする(ステップS27)。こうして、ディスクキャッシュメモリ上へのデータのライトが終了した後、そのブロックをLRU最上位に位置付けてCPUにライト終了を通知する

(ステップS28、S29)。方式によっては、最下位ブロックがダーティブロックのとき最下位に最も近いダーティでないブロックを選択してそれを割当てるなど異なる動作を行なう場合もあるが、ここでは前述した図示の従来例に沿って説明していく。

【0009】図6は、従来のライト時の管理テーブルの操作例の説明図である。この図は、新たに割当てたブロックにデータをライトした場合のキャッシュ管理テーブル上の操作の例を示したものである。この場合、LRU最上位を示すポインタを、その新たなブロックに対応する管理ブロック45を指し示すように変更する。そして、管理ブロック45のLRUポインタは、以前LRU最上位であった管理ブロックを指し示すように設定する。既にLRUリンクに繋がれているブロックをLRU最上位に位置付ける場合は、まずその前後の管理ブロックのLRUポインタを操作してリンクから外した後に、同様の操作を行なう。

【0010】図7及び図8は、ともに従来のライトバックの実行例を示す図である。これらの図は、従来のライトバックの実行タイミングの例を示したものであり、どちらの図も横軸は右方向への時間軸を示し、太線の部分はディスクキャッシュメモリ上にダーティブロックが1ブロック以上存在する期間を示す。図7は、ライト直後にライトバックを行なう場合であり、例えば、ライト

(1)の動作終了後すぐにそのデータをライトバックする(ライトバック(1))というように実行する。それに対して図8は、一定時間T(通常数秒〜数分)毎にその時点でダーティであったブロックを連続して書き出す方式である。この他に、ダーティブロックがLRUで追い出されたときにだけライトバックを行なう方式や、CPUからのリード/ライト要求の時間間隔を監視して一定時間以上ライト要求がなかったら、その後しばらくリード/ライト要求がないものとみなしてその間に集中してライトバックを行なう方式がある。また、それらの組合せも考えられている。

【0011】

【発明が解決しようとする課題】しかしながら、上述した従来の技術には、次のような問題があった。即ち、ライトバック方式のディスクキャッシュ装置において、ライトバックの動作は磁気ディスク装置のアクセス時間によって制限されるためにライト動作に比べて大幅に遅い。従って、短時間のうちに連続してライト要求を受けた場合はディスクキャッシュメモリ上のダーティブロックの数が増え、最後にはディスクキャッシュメモリ上のブロックすべてがダーティブロックになって、以後ライトバックが間に合わない間はライト要求の度に磁気ディスク装置をアクセスしなければならなくなる。また、リードミスが発生した場合は磁気ディスク装置をアクセスする必要があるため、ライトバック動作と競合してアクセス時間が長くなる確率も高くなる。ライトアクセスの

頻度が高い使用環境ではこれらがネックとなって、スループットが低下する。

【0012】これを防ぐには、ライトバック時にヘッドスケジューリングを行なって磁気ディスク装置のシーク時間を短縮したり、連続する複数のブロックをまとめて1回のライトバックで終了させてスループットを向上させることが必要である。しかし、前述の管理方式では、ディスクキャッシュメモリ上のブロックはLRU順位以外はそれぞればらばらに管理される。このため、ヘッドスケジューリングや連続ブロックの探索を行なうにはライトバック時に毎回キャッシュ管理テーブル全域を検索する必要があり、オーバヘッドが極端に大きくなるため現実的ではない。また、ブロック管理情報としてLRUポインタとは別のポインタを用意して全ブロックを磁気ディスク装置のアドレス順に繋げるリンクを張ってライトバック時にこれを参照するという手段も考えられる。しかしこの場合、リード/ライト時に新たなブロックを割り当てる度にキャッシュ管理テーブルを検索してアドレス順のリンクに繋げる必要があるが、磁気ディスク装置のアドレス空間は広大であり、このようなアドレス順のリンクを保守するのにもかなりのオーバヘッドがかかるため、これも効率的とはいえない。

【0013】また、ライトバックの実行タイミングもスループットに大きな影響を及ぼす。例えば、ライトの直後にライトバック動作を起動する図7の方式ではライト回数とほぼ同じ回数のライトバックを行なうことになるため、スループットを上げられない。その上、その動作時間の差により図中のライト(2)以降のようにライトが連続すると、ライトバックが追いつかず、ディスクキャッシュメモリの大半がダーティである状態に陥りやすい。図8では、Tの時間内に同一ブロックへのライトが複数回発生すればその分ライトバック回数を減らすことができるが、時間T後にはどのブロックも一律に書き出してしまふ。このため、Tが短ければその効果は小さくなり、長ければ大半のブロックがダーティとなる可能性が大きくなる。そして、この場合では短時間にライトバックを集中して行なうためにこの間は極端にスループットが低下するという問題がある。その他の方式でも同様に、ライト回数に対してライトバック回数を大幅に減らすことは期待できず、スループットをあまり上げることができない。

【0014】本発明は、以上の点に着目してなされたもので、ライトアクセスの頻度が高い使用環境ではスループットが低下するという問題を回避するために、(1)僅かなオーバヘッドで1トラック分程度の範囲でのヘッドスケジューリングを可能にして磁気ディスク装置のシーク時間を短縮するとともに、複数の連続するブロックをまとめて書き込んでライトバックの回数を減らすことを可能として、スループットの低下を抑えることができるライトバック式ディスクキャッシュ装置を提供するこ

とを目的とするものである。

【0015】

【課題を解決するための手段】本発明のライトバック式ディスクキャッシュ装置の第1の発明は、ディスク装置上に格納されたデータをブロック毎に転記するディスクキャッシュメモリと、前記ディスク装置へのデータの書き込みの発生時は、当該ディスクキャッシュメモリに転記されている該当するブロック内のデータのみを更新するデータ書き込み手段と、前記ディスク装置上の連続する1トラック分程度の複数のブロックを1つのグループとし、当該グループ内のブロックを前記ディスク装置上の同一トラック内のアドレス順にリンクして管理するキャッシュ管理テーブルと、当該キャッシュ管理テーブル上に前記ディスクキャッシュメモリ上でのみ更新されているブロックを含むグループがあるときは、当該グループ内の連続する複数のブロックを1回のアクセスにより前記ディスク装置上に書き戻すライトバック手段とから成ることを特徴とするものである。

【0016】本発明のライトバック式ディスクキャッシュ装置の第2の発明は、ディスク装置上に格納されたデータをブロック毎に転記するディスクキャッシュメモリと、前記ディスク装置へのデータの書き込みの発生時は、当該ディスクキャッシュメモリに転記されているブロック内のデータのみを更新するデータ書き込み手段と、前記キャッシュメモリ上でのみ更新されているブロックを前記ディスク装置上に書き戻すライトバック手段と、当該ライトバック手段の動作が発生する時間間隔を決定するタイマ値を、前記ディスクキャッシュメモリ上でのみ更新されているブロックの数が多くなるに従って小さく設定するタイマ値設定手段とから成ることを特徴とするものである。

【0017】

【作用】本発明のライトバック式ディスクキャッシュ装置には、第1及び第2の発明に対応して以下の2つの作用がある。

(1) キャッシュ管理テーブルによりディスク装置上の連続する1トラック分程度の複数ブロックにわたる領域の一括管理が可能となる。このために、キャッシュ管理テーブル上の管理ブロック内に個々のブロックをディスク装置上の領域内でのアドレス順にリンクするためのリンクポインタを設ける。そして、リンクされたブロック群にダーティブロックが含まれる場合はグループ内の先頭ブロックのLRUポインタのみを有効とする。これにより、このグループ全体をLRU管理上の1単位とするキャッシュ管理方式を設けて、ライトバック時にはダーティブロックを含むグループの内でもLRUで最も下位に属するものの中で、下位アドレスのブロックから順にできる限り複数ブロックを一括して書き出すようにする。

(2) ライト又はライトバックの動作終了後、次にライトバックを起動するタイミングが可変のタイマ(ライト

バックタイマ)により決定される。そのタイマ値はその時点でディスクキャッシュメモリ上に残っているダーティブロックの割合によって決定される。

【0018】

【実施例】以下、本発明の実施例を図面を参照して詳細に説明する。図1は、本発明のライトバック式ディスクキャッシュ装置の一実施例のブロック図である。キャッシュ管理テーブル1は、RAMから成り、ディスクキャッシュメモリ2内の各キャッシュブロック4に対応する管理ブロック3を格納している。ディスクキャッシュメモリ2は、RAMから成り、ディスク装置10に格納されたブロック毎のデータをキャッシュブロック4として転記する。ディスクキャッシュ制御部6は、データ書き込み手段7と、ライトバック手段8と、タイマ値設定手段9とから成る。データ書き込み手段7は、CPUからディスク装置10へのデータの書き込みの発生時は、ディスクキャッシュメモリ2に該当するブロックが転記されている場合には、そのブロック内のデータのみを更新する。

【0019】ライトバック手段8は、キャッシュメモリ2上でのみ更新されているブロックをディスク装置10上に書き戻す。より具体的には、ライトバック手段8は、キャッシュ管理テーブル1によりディスクキャッシュメモリ2上でのみ更新されているブロックを含むグループがあるときは、当該グループ内の連続する複数のブロックを1回のアクセスによりディスク装置10上に書き戻す。ライトバック手段8の動作は、タイマ値設定手段9により決定されるタイマ値で決定される時間間隔で発生する割込み信号により起動される。タイマ値設定手段9は、タイマ割込みのためのタイマ値を、ディスクキャッシュメモリ2上でのみ更新されているブロックの数が多くなるに従って小さい値に設定する。即ち、ディスクキャッシュメモリ2上のキャッシュブロックの総数に対するダーティブロックの割合が高くなる程タイマ値を小さくする。ダーティブロックの計数は、キャッシュ管理テーブル1等にダーティブロック総数を情報として持ち、これを参照する方法により行なうことができる。

【0020】管理ブロック3は、ディスクキャッシュメモリ2上のブロック4と一体一で対応しており、ここにはブロック内に格納されているデータの磁気ディスク装置10上でのアドレスとそのブロックがダーティであるかどうかを示すダーティフラグ、LRU順位を管理するためのLRUポインタ、及びグループ間リンクポインタなどの情報が含まれている。また、ディスクキャッシュメモリ2の図上特に他のブロックと区別したブロックはダーティブロックであることを示している。そして、ブロックに格納されるデータのコピー元が、磁気ディスク装置上で連続する数ブロック〜数十ブロック分の大きさの同じ領域に含まれるブロックと呼び、これらのブロックの間をアドレス順にリンクするためのポインタがグル

ープ間リンクポインタである。ブロックサイズ及び1グループ分の領域は任意であるが、例えば、それぞれ4kB、64kBであるとする、1グループの最大ブロックは16である。図示の例ではこの1グループ分の領域は磁気ディスク装置の1トラックに対応するものとして説明する。

【0021】図1で、a0、a1、b0等はブロック管理情報ブロックであり、それぞれにディスクキャッシュメモリ上のブロックA0、A1、B0等が対応する。そして、1つのグループに含まれるブロック群はAn、Bnのように1文字目で区別され、それに続く数字はグループ内でのそのブロックのアドレスを示すものとする。そして、図中ではそれぞれのグループ間リンクによってアドレス順にリンクされている様子を横向き矢印で表わしている。このリンクはd7からd9のように不連続なブロックの間もリンクすることが可能である。また、aとdのグループはダーティブロックを含むため、グループ単位でLRUリンクに繋がられている。これはグループ内で先頭に位置するブロックの管理情報内のLRUポインタのみを有効とし、他のブロックのものは無効として、先頭ブロックのLRU順位をグループ全体の順位とみなすことで実現することができる。また、グループが1つのブロックだけから成る場合はグループ間リンクポインタを無効にしておけばよい。

【0022】図1において、次のライトバックの対象となるのは、Dグループであり、D3からD7は連続したブロックであるためまずこのブロック群を一括して書き出すことができ、これによりライトバック回数を削減することができる。なお、D4からD6のブロックのいくつかはダーティではないがディスクキャッシュメモリ上に存在するというような場合は、それらも含めて一度で書き出す場合の磁気ディスク装置のアクセス時間及びデータ転送時間と、複数回に分けて書き出す場合のそれを推定して比較し、総合的に最も短い時間でライトバックできるパターンを選択するものとする。そして、さらにその次のライトバックではブロックD9を書き出すというように選択する。このように、同一トラック上のデータをアドレス順に書き出すことで局所的なヘッドスケジューリングを行なうことができ、磁気ディスク装置のアクセス時間を短縮することも可能になる。

【0023】CPUからのディスクアクセスの局所性により、このように1トラック分ほどの領域に複数のブロックが含まれることは多いものと期待できる。さらに、このグループ単位でLRUの管理を行なうため、例えば小容量の一時ファイルへのライトのような比較的狭い領域への書き込みや、データベースの更新履歴情報のライトのような小容量のデータのシーケンシャルな書き込みが頻繁に要求されるような場合は特に、該当するグループはLRUの上位に長くどまりつつ、その間に次々と新たなブロックを連結していく。これにより、この領域

へのライト要求はディスクキャッシュメモリへのライトだけで終了させることができる。そして、この領域へのライト要求がなくなるとLRUの順位が下がり、ライトバックの対象となった際に、一括して書き出すことによりライト回数に対するライトバック回数の大幅な削減が可能となる。この結果、スループットを大幅に向上させることができる。

【００２４】図９は、本発明によるライト時のキャッシュ制御を説明するフローチャートである。この図は、ディスクキャッシュ装置がCPUからのライト要求を受けたときの、本発明の管理方式におけるキャッシュ制御の例をフローチャートで示したものである。ディスクキャッシュ装置は、まずキャッシュ管理テーブルを検索して該当領域を含むブロックの有無を判定する（ステップS1、S2）。本発明では、この検索においては同一のトラックに含まれるブロック群を同時に検索する必要があるため、磁気ディスク装置の装置番号、シリンダ番号、及びヘッド番号までが一致するものをキャッシュ管理テーブル上の管理ブロックを参照して探索していく。そして、一致するものを発見したらグループ間リンクを前後に手繰って該当するブロックを探せばよい。なお、このようにすると、単に該当するブロックのみを探すよりも高速に検索できる可能性が高い。

【００２５】このように検索して、該当するブロックがディスクキャッシュメモリ上に存在した場合はそのブロックにデータを上書きした後にそのブロックを含むグループをLRUの最上位に位置付け直す（ステップS13、S11）。そうでなかった場合は、従来の方法と同様に必要ならば（ステップS3、S4）、LRU最下位のブロックをライトバックして（ステップS5）、空きブロックを割り当てて（ステップS6）その領域にデータをライトする（ステップS7）。そして、空きブロックを割り当てた場合で、同一トラックのデータを格納したブロックが他に存在する場合は（ステップS8）、そのグループ間リンクのアドレス順で該当する順位にその新たなブロックを挿入し（ステップS9）、他になければ単独で1つの新規グループとして（ステップS10）、それぞれのグループ全体をLRU最上位に位置付け（ステップS11）、CPUにライト終了を通知する（ステップS12）。このとき、既存のグループがダーティブロックを含まなかった場合は各ブロックごとにLRUリンクが張られているわけであるが、この場合はそれらのブロック全てをLRUリンクから外して新たに1つのグループとしてLRUの最上位に位置付ける。それ以外に、ダーティブロックで挟まれたブロック群、あるいは連続するダーティなブロック群のみをひとまとめのLRU管理単位とする方式も考えられる。このグループ間リンクは数ブロックから数十ブロックをアドレス順にリンクするだけのものである。これを保守するためのオーバーヘッドは無視できる程度のものであり、前述のよ

うにキャッシュメモリテーブルの検索を高速化させる効果さえもある。こうして、ディスクキャッシュメモリへのデータの格納、ブロック管理情報の更新を終了したところでCPUにライト終了を通知するのである。

【００２６】図１０は、本発明によるライト時の管理テーブルの操作例の説明図である。ここで、新たに割り当てたブロックにデータをライトした場合、本発明によるキャッシュ管理テーブル上の操作の例を図１０に沿って説明する。この図では、b2で示されるブロックを新たにリンクするに当たって、既にb1、b3及びb7で示される同一トラック内のブロック群が1つのグループとしてリンクされていたことを示している。この場合、まず既存のグループをアドレス順に手繰ってb2をb1とb3の間のグループ間リンクに挿入する。その上で、このグループ全体を以前のLRUリンクから外してLRUの最上位に位置付けるようにLRUポインタとLRU最上位ポインタを操作するのである。

【００２７】次に、本発明によるライトバックのタイミング制御方式を説明する。本発明では、ライト又はライトバックの動作終了毎に、ディスクキャッシュメモリ上に占めるダーティブロックの割合を判定し、その割合に応じて次にライトバックを起動するためのタイマ（以後ライトバックタイマとよぶ）のタイマ値を設定して一定時間毎のカウンタを開始する。そして、設定されたタイマ値分のカウンタを終了した時点で次のライトバックを起動することにより、ライトバックのタイミングを制御するのである。

【００２８】図１１は、本発明によるライトバック間隔決定を説明するフローチャートである。この図は、ライトバックタイマのタイマ値を決定するフローチャートの例を示したものである。この例では、ライト終了時にすでにライトバックタイマが起動されていなければ（ステップS31）タイマ値をT1として起動する（ステップS32）。そして、ライトバック終了時点では、ディスクキャッシュメモリ上のダーティブロックの割合に応じて（ステップS33）、1/2まではT2（ステップS34）、2/3まではT3（ステップS35）、3/4まではT4とし（ステップS36）、それ以上の時は即座に次のライトバックを起動する（ステップS37）。この場合T1はT2より大きい。等しく、 $T2 > T3 > T4$ であり、それらは実際には数百ミリ秒から数分程度の値が適当である。

【００２９】図１２は、本発明によるライトバックの実行例を示す図である。この図は、本発明によるライトバックの実行タイミングの例を示したものであり、横軸は右方向への時間軸を示している。この図では、ライトバック（0）を終了した時点でディスクキャッシュメモリ上にダーティブロックが存在しなくなったことを示している。すると、次にライトされるまではライトバックタイマは停止状態になる。次にライト（1）の要求を受け

たとすると、動作終了後にはT1なるタイマ値でライトバックタイマを起動する。そして、T1の時間後にライトバックを行ない、その終了時点でディスクキャッシュメモリ上のダーティブロックの割合が1/2以下であったとすると次はT2なるタイマ値で起動する。こうしてライトバックを実行していき、ライトバック(3)で再びダーティブロックがなくなった時点でライトバックタイマを停止する。

【0030】このようにタイミングを制御することで、ダーティブロックの割合が少ない間は少しずつ、割合が大きくなるほど短い間隔でLRU順位で下位のブロックからライトバックすることになる。このことにより、ライト直後に即ライトバックを行なったり、一定期間ごとにすべてのダーティブロックのライトバックを行なうような方式よりも、LRUで上位のブロックほど、ライトバックの順位が低くその上ダーティブロックの割合が少なくなるとそのブロックをライトバックする順番が近づくほどライトバックの間隔が長くなるために、ディスクキャッシュメモリ上に長く留まるようになる。また、ダーティブロックの割合が増えるほど頻繁にライトバックを行なうようになるため、キャッシュメモリの大半がダーティブロックとなるような状況にはなりにくくなる。

【0031】

【発明の効果】以上説明したように、本発明のライトバック式ディスクキャッシュ装置によれば、(1)1トラック分程度の領域に含まれるブロックをグループとしてグループ間リンクポインタによるリンクでアドレス順に管理し、ダーティブロックを含む場合はそのグループをLRU管理上の単位とした上で、ライトバック時にはそのグループ間リンクを参照することにより、その領域内の可能な限りの複数ブロックの一括書き出しを行なうことによりライトバック回数を削減し、かつ、磁気ディスク装置のヘッドのシーク動作が不要な同一トラック内でアドレス順に書き出していくことによって、局所的なヘッドスケジューリングによる磁気ディスク装置のアクセス時間の短縮を僅かなオーバーヘッドで容易に実現するようにしたので、ライトアクセスの頻度高い使用環境においてもスループットを従来以上に高めることが可能となる。さらに、グループ単位でLRUの管理を行なうため、例えば小容量の一時ファイルへのライトのような比較的狭い領域への書き込みや、データベースの更新履歴情報のライトのような小容量のデータのシーケンシャルな書き込みが頻繁に要求されるような場合は特に、該当するグループはLRUの上位に長く留まりつつ、その間に次々と新たなブロックを連結していくことで、この領域へのライト要求はディスクキャッシュメモリへのライトだけで終了させることができる。そして、この領域へのライト要求がなくなってLRUの順位が下がりライトバックの対象となった際に、一括して書き出すことによりライト回数に対するライトバック回数の大幅な削減が

可能となり、スループットを大幅に向上させることができる。また、(2)ライト/ライトバック終了後、次にライトバックを起動するまでの時間を、ディスクキャッシュメモリ上にダーティブロックが占める割合が少ないほど長く、多いほど短く、そして一定以上の割合だったら間隔をあけずに連続して行なうというように、ダーティブロックの割合によって調整するようにしたので、LRUで上位のブロックほど、ライトバックの順位が低い上にダーティブロックの割合が少なくなるとそのブロックをライトバックする順番が近づくほどライトバックの間隔が長くなる。従って、ディスクキャッシュメモリ上に従来以上に長く留まるようになる。このため、同一ブロックへのライトが頻繁に発生する場合、そのブロックはディスクキャッシュメモリ上に長く留まって、ライト回数に対してライトバックの回数を従来以上に削減できるようになる。このため、スループットを向上させることができる。そして、このライトバックタイミングの制御方式を上述した本発明(1)によるディスクキャッシュメモリ管理方式と組み合わせた場合、特に前述のような狭い領域への書き込みや小容量のデータのシーケンシャルな書き込みが頻繁に要求された場合、その領域を含むグループは従来のライトバックタイミングの制御方式より一層長くディスクキャッシュメモリ上に留まるため、ライトバック回数の更なる削減によるスループットの一層の向上が期待できる。また、ダーティブロックの割合が増えるほど頻繁にライトバックを行なうようになるため、キャッシュメモリの大半がダーティとなるような状況にはなりにくく、その上で適当な時間間隔でライトバックを行なう。このため、一定時間ごとに集中してライトバックを行なう場合のように一時的にスループットが極端に低下するようなことも発生しないという利点もある。

【図面の簡単な説明】

【図1】本発明のライトバック式ディスクキャッシュ装置の一実施例のブロック図である。

【図2】従来のディスクキャッシュ装置のリード時の動作を説明するブロック図である。

【図3】従来のディスクキャッシュ装置のライト時の動作を説明するブロック図である。

【図4】従来のライトバック式ディスクキャッシュ装置の管理方式の説明図である。

【図5】従来のライト時のキャッシュ制御を説明するフローチャートである。

【図6】従来のライト時の管理テーブルの操作例の説明図である。

【図7】従来のライトバックの実行例を示す図である。

【図8】従来のライトバックの実行例を示す図である。

【図9】本発明によるライト時のキャッシュ制御を説明するフローチャートである。

【図10】本発明によるライト時の管理テーブルの操作

10

20

30

40

50

例の説明図である。

【図11】本発明によるライトバック間隔を決定する例を説明するフローチャートである。

【図12】本発明によるライトバックの実行例を示す図である。

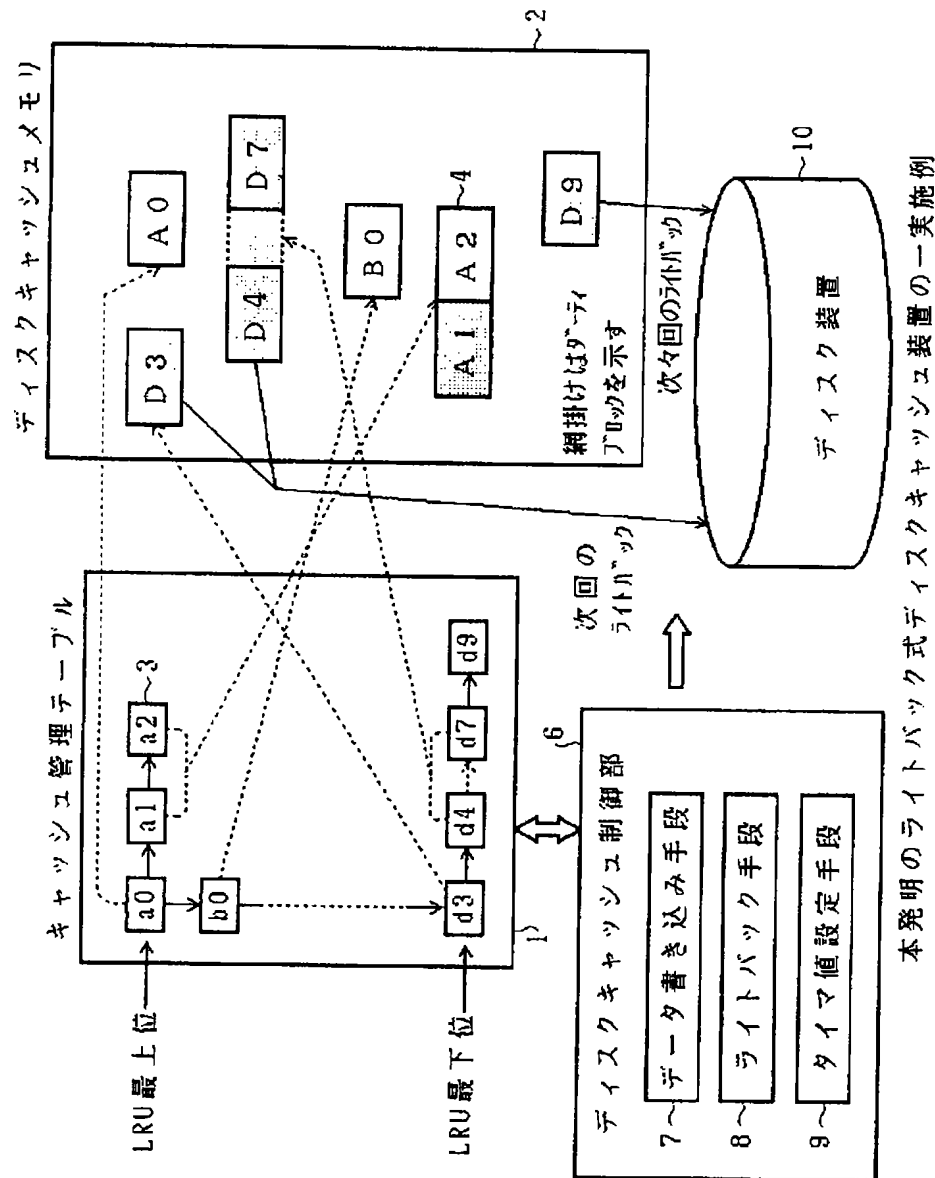
【符号の説明】

- 1 キャッシュ管理テーブル
2 ディスクキャッシュメモリ

- * 3 管理ブロック
4 キャッシュブロック
6 ディスクキャッシュ制御部
7 データ書き込み手段
8 ライトバック手段
9 タイマ値設定手段
10 ディスク装置

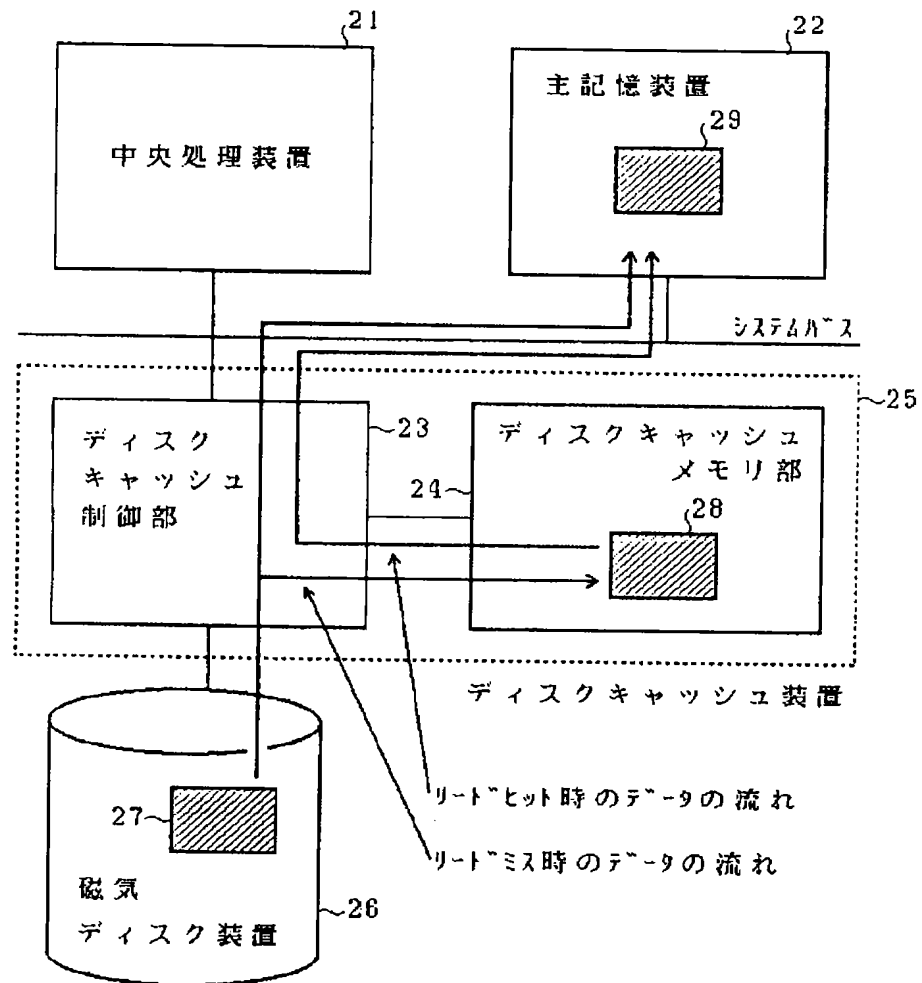
*

【図1】



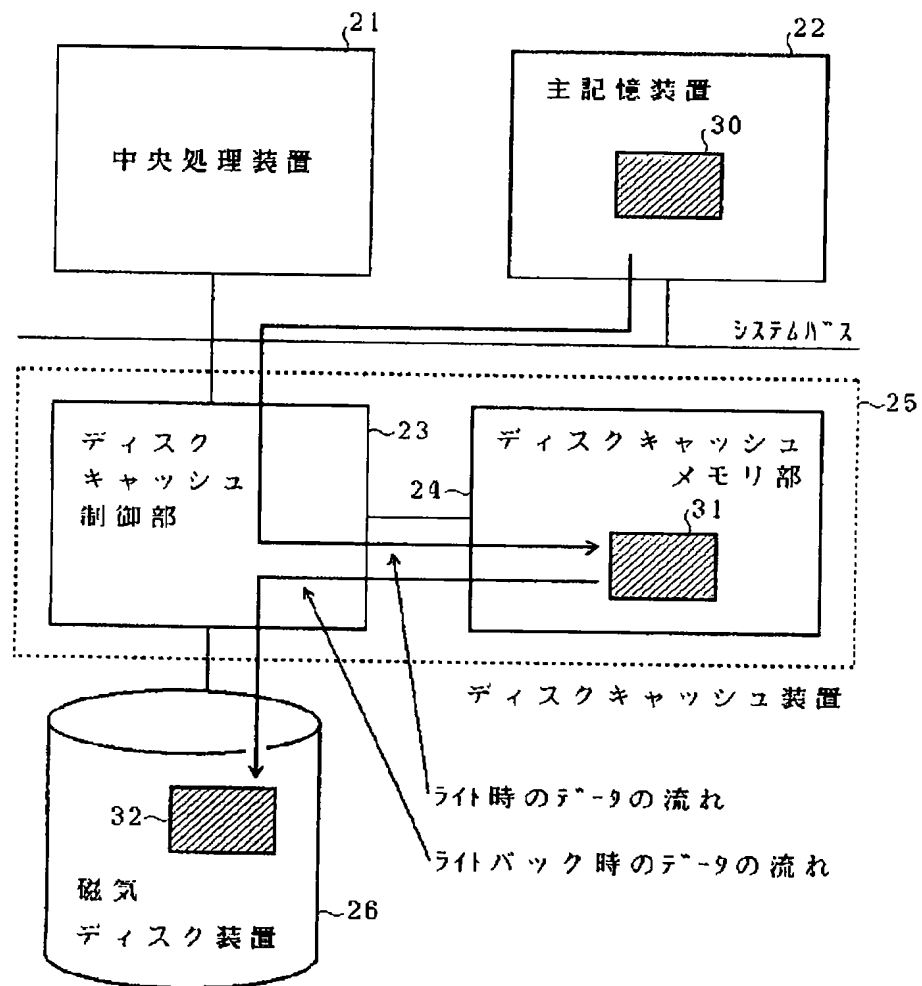
本発明のライトバック式ディスクキャッシュ装置の一実施例

【図2】



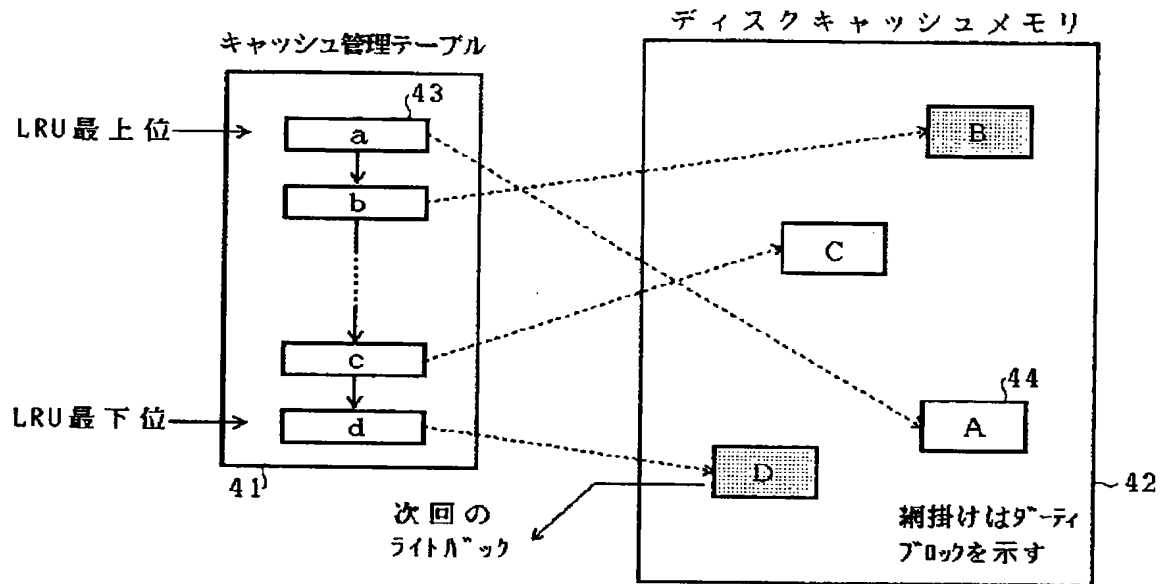
リード時のディスクキャッシュ装置の動作

【図3】



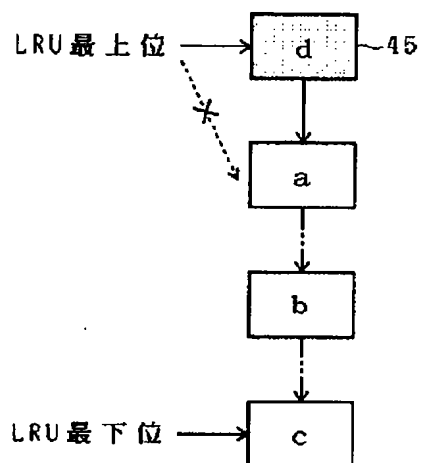
ライト時のディスクキャッシュ装置の動作

【図4】



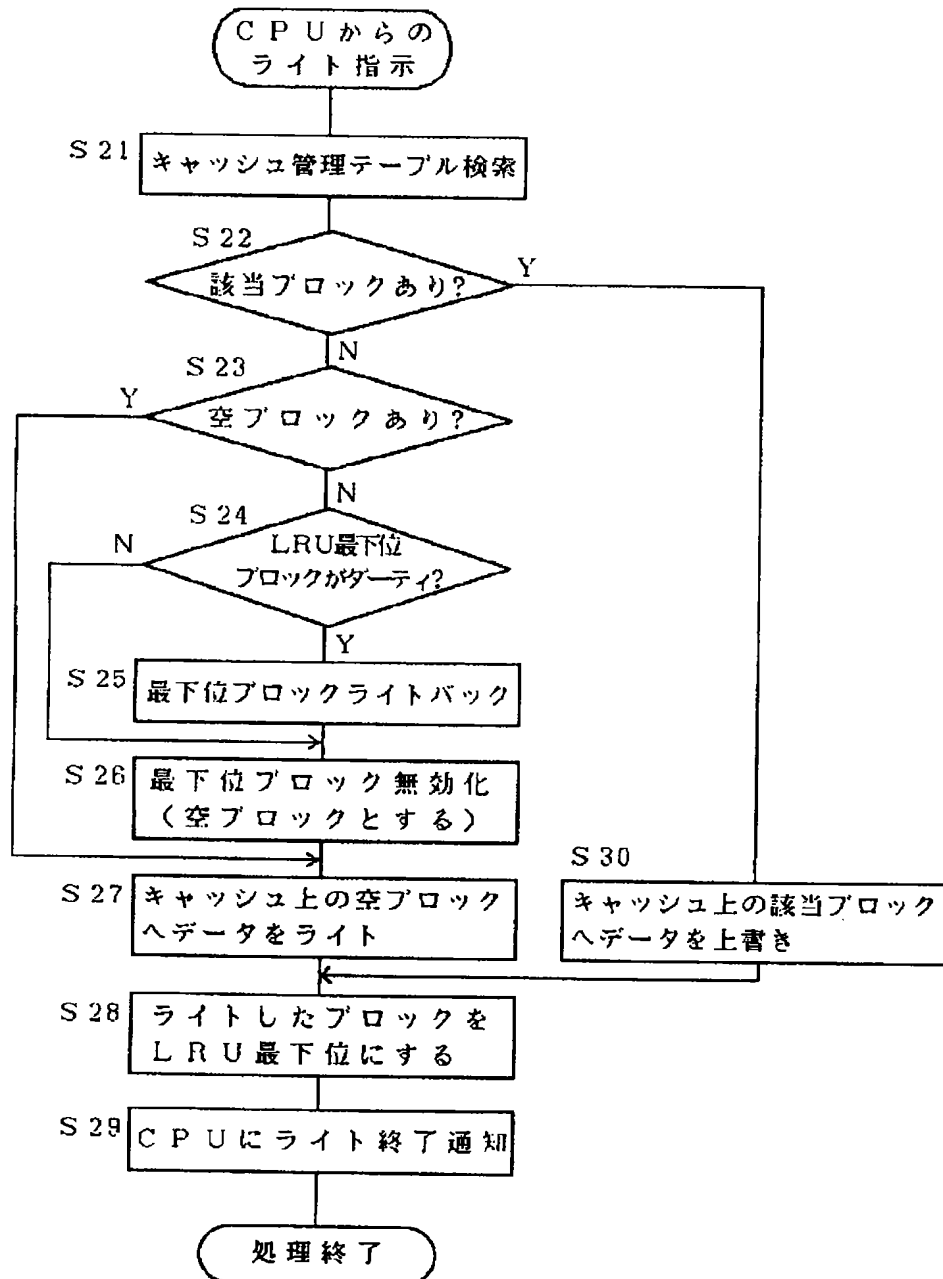
従来のライトバックキャッシュの管理方式

【図6】



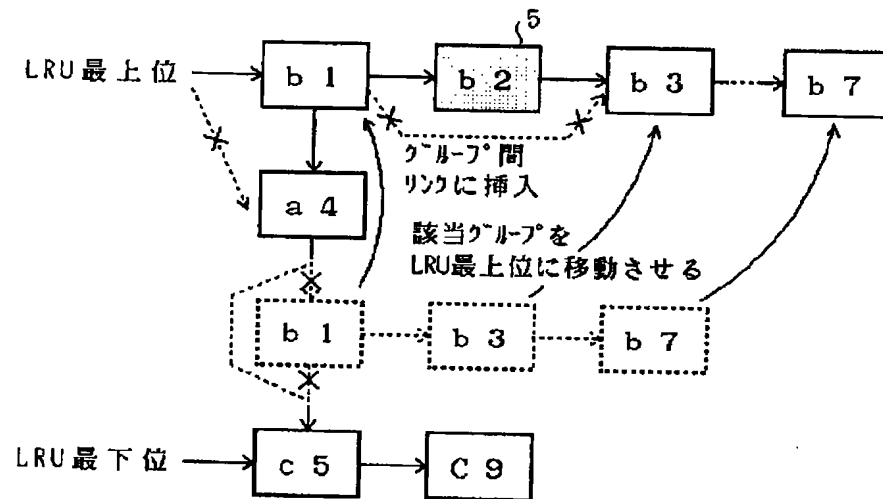
従来のライト時の管理テーブルの操作例

【図5】



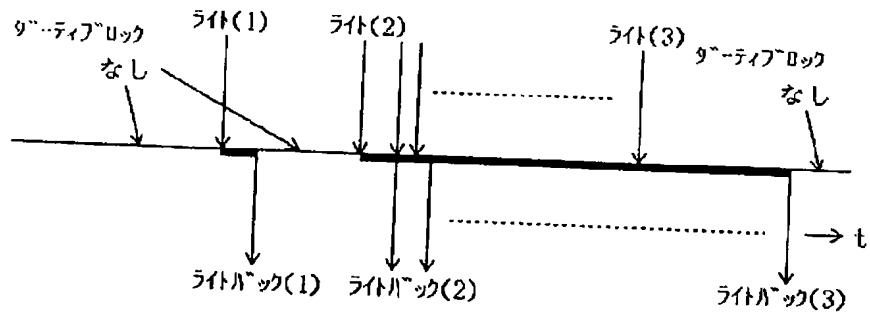
従来のライト時のキャッシュ制御

【図10】



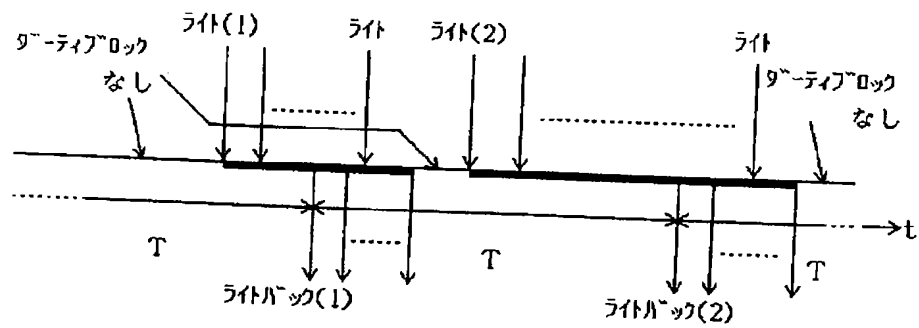
本発明によるライト時の管理テーブルの操作例

【図7】



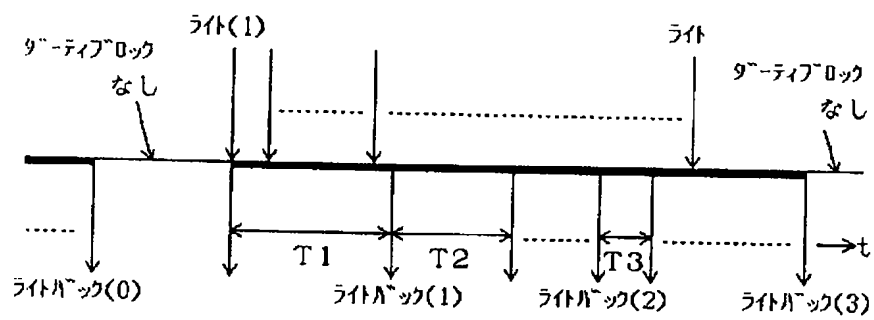
従来のライトバックの実行例

【図8】



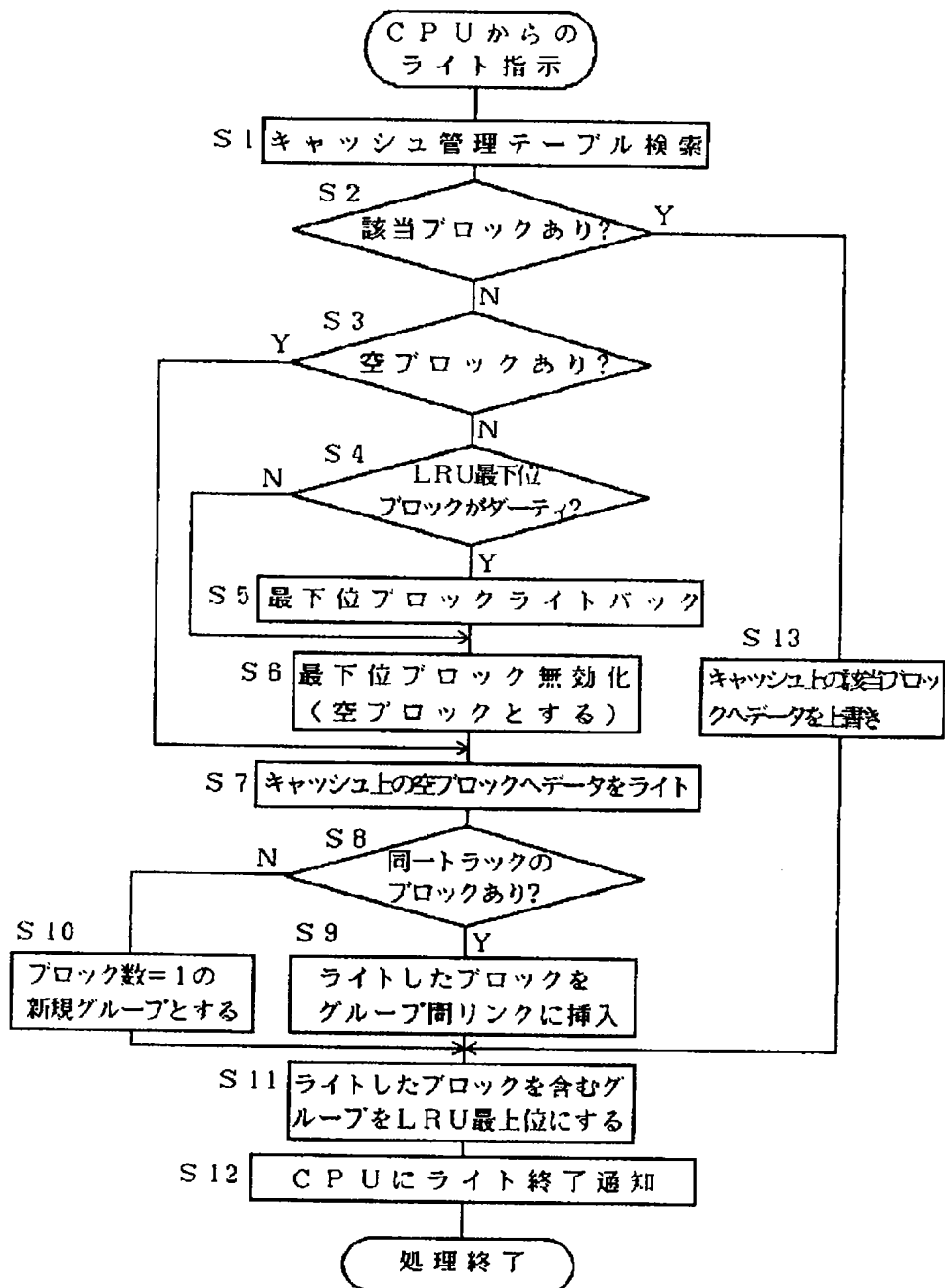
従来のライトバックの実行例

【図12】



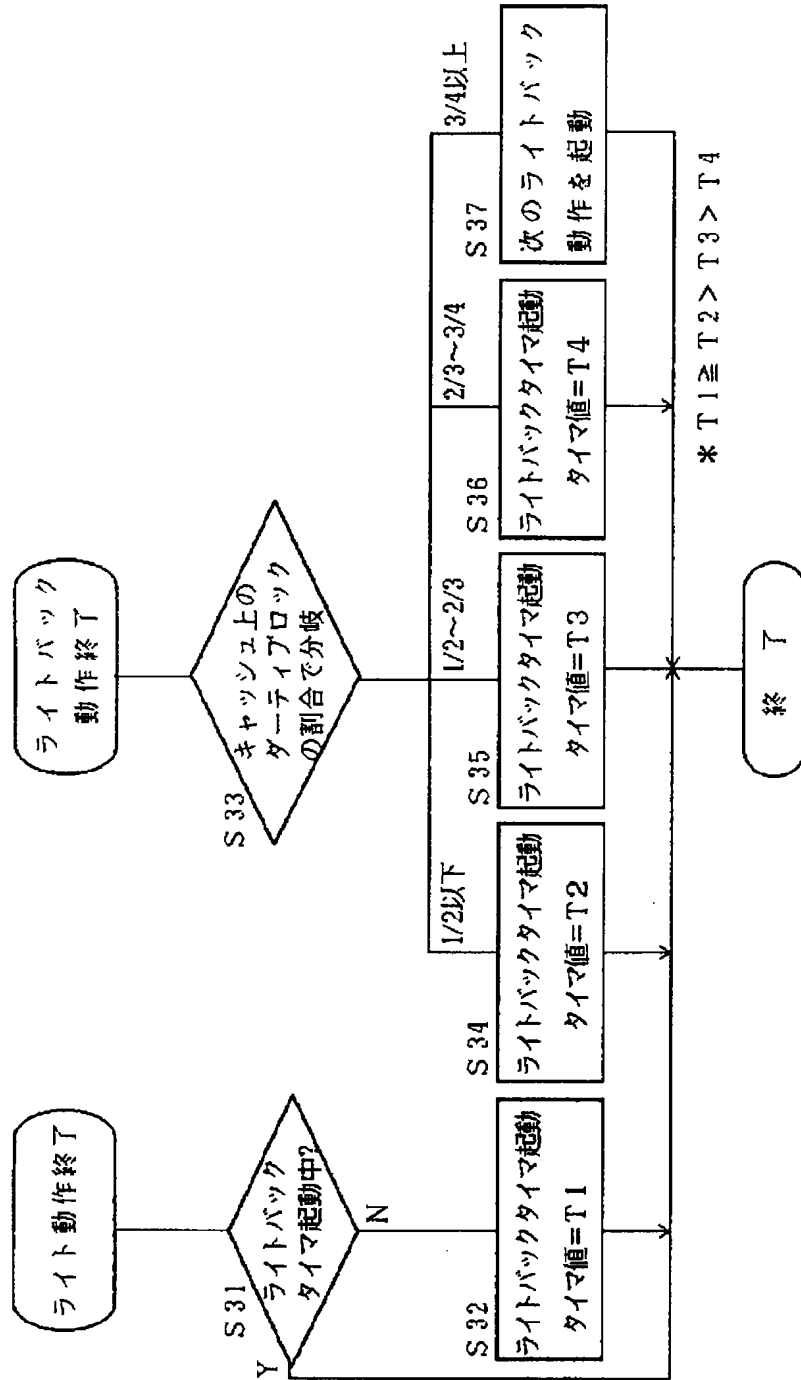
本発明によるライトバックの実行例

【図9】



本発明によるライト時のキャッシュ制御

【図11】



本発明によるライトバック間隔決定例